PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference	FOR FURTHER ACTION	see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.
BB1246		
International application No.	International filing date *(day/month/year)*	(Earliest) Priority Date *(day/month/year)*
PCT/US 99/20420	07/09/1999	08/09/1998
Applicant		
E. I. DU PONT DE NEMOURS AND COMPANY et al.		

This International Search Report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This International Search Report consists of a total of 6 sheets.

[X] It is also accompanied by a copy of each prior art document cited in this report.

1. **Basis of the report**

   a. With regard to the **language**, the international search was carried out on the basis of the international application in the language in which it was filed, unless otherwise indicated under this item.

   [ ] the international search was carried out on the basis of a translation of the international application furnished to this Authority (Rule 23.1(b)).

   b. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international search was carried out on the basis of the sequence listing :

   [X] contained in the international application in written form.

   [X] filed together with the international application in computer readable form.

   [ ] furnished subsequently to this Authority in written form.

   [ ] furnished subsequently to this Authority in computer readble form.

   [ ] the statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.

   [ ] the statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished

2. [ ] **Certain claims were found unsearchable** (See Box I).

3. [X] **Unity of invention is lacking** (see Box II).

4. With regard to the **title**,

   [X] the text is approved as submitted by the applicant.

   [ ] the text has been established by this Authority to read as follows:

5. With regard to the **abstract**,

   [X] the text is approved as submitted by the applicant.

   [ ] the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. The figure of the **drawings** to be published with the abstract is Figure No. _____

   [ ] as suggested by the applicant.

   [ ] because the applicant failed to suggest a figure.

   [ ] because this figure better characterizes the invention.

   [X] None of the figures.

INTERN NAL SEARCH REPORT

International application No.
PCT/US 99/20420

B x I Obs rvations wher ertain claims wer f und uns archable (Continuati n of it m 1 f first heet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. [X] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

[X] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 23 completely; 1-22 partially

   Isolated nucleotide sequence that encodes a thioredoxin polypeptide from Momordica charantia, represented by SEQID 1 and 2, respectively; furthermore; the recombinant expression of the same, a method for selecting a thioredoxin sequence that affects thioredoxin expression in a plant cell and a method for isolating homologous sequences.

2. Claims: 24 completely; 1-22 partially

   Isolated nucleotide sequence that encodes a thioredoxin polypeptide from Catalpa speciosa, represented by SEQID 3 and 4, respectively; furthermore; the recombinant expression of the same, a method for selecting a thioredoxin sequence that affects thioredoxin expression in a plant cell and a method for isolating homologous sequences.

3. Claims: 25,26 completely; 1-22 partially

   Isolated nucleotide sequences that encode thioredoxin polypeptides from soybean, represented by SEQID 5,7 and 6,8, respectively; furthermore; the recombinant expression of the same, a method for selecting a thioredoxin sequence that affects thioredoxin expression in a plant cell and a method for isolating homologous sequences.

4. Claims: 27 completely; 1-22 partially

   Isolated nucleotide sequence that encodes a thioredoxin polypeptide from Vernonia, represented by SEQID 9 and 10, respectively; furthermore; the recombinant expression of the same, a method for selecting a thioredoxin sequence that affects thioredoxin expression in a plant cell and a method for isolating homologous sequences.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/20420

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/29 C12N15/53 C12N15/82 C12Q1/68 C07K14/415
C12N9/02

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C12Q A01H C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	RIVERA-MADRID, R., ET AL.: "Nucleotide sequence of a cDNA clone encoding an Arabidopsis thaliana thioredoxin h" PLANT PHYSIOL., vol. 102, 1993, pages 327-328, XP002131240 the whole document ---	1-27
A	RIVERA-MADRID,R., ET AL.: "evidence for five divergent thioredoxin h sequences in Arabidopsis thaliana" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE USA, vol. 92, June 1995 (1995-06), pages 5620-5624, XP002131241 page 5621; figure 1 ---	1-27

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
19 May 2000	02.06.00

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Holtorf, S

3

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	MARTY, I., ET AL.: "Nucleotide sequence of a cDNA encoding a tobacco thioredoxin" PLANT MOLECULAR BIOLOGY, vol. 17, pages 143-147, XP002131242 the whole document	1-27
A	WO 96 03505 A (AGRONOMIQUE INST NAT RECH ;GAUTHIER MARIE FRANCOISE (FR); LULLIEN) 8 February 1996 (1996-02-08) the whole document	1-27
A	SZEDERKENYI, J., ET AL.: "cDNA expressed in Ricinus cotyledons" EMBL SEQUENCE DATA LIBRARY, 4 April 1996 (1996-04-04), XP002131252 heidelberg, germany accession no.Z70677	1-27
A	WO 95 19443 A (CIBA GEIGY AG ;WARD ERIC R (CH); ALEXANDER DANNY C (US); UKNES SCO) 20 July 1995 (1995-07-20) see SEQID 13 page 12	1-27
A	BRUGIDOU, C., ET AL.: "the Nicotiana tabacum genome encodes two cytoplasmic thioredoxin genes which are differently expressed" MOLECULAR AND GENERAL GENETICS,1993, pages 285-293, XP002129274 the whole document	1-27
A	WO 93 08274 A (REGENTS OF THE UNIVERSITY OF CALIFORNIA) 29 April 1993 (1993-04-29) the whole document	1-27
A	SHI JINRUI ET AL: "A novel plasma membrane-bound thioredoxin from soybean." PLANT MOLECULAR BIOLOGY 1996, vol. 32, no. 4, 1996, pages 653-662, XP002138131 ISSN: 0167-4412	1-27
A	BUCHANAN B B ET AL: "Thioredoxin: A multifunctional regulatory protein with a bright future in technology and medicine" ARCHIVES OF BIOCHEMISTRY AND BIOPHYSICS,US,NEW YORK, US, vol. 314, no. 2, 1 November 1994 (1994-11-01), pages 257-260, XP002089742 ISSN: 0003-9861 the whole document	1-27

-/--

3

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	WO 96 12799 A (UNIV CALIFORNIA) 2 May 1996 (1996-05-02) the whole document	1-27
A	GAUTIER MARIE-FRANCOISE ET AL: "Characterization of wheat thioredoxin h cDNA and production of an active Triticum aestivum protein in Escherichia coli." EUROPEAN JOURNAL OF BIOCHEMISTRY MARCH 1, 1998, vol. 252, no. 2, 1 March 1998 (1998-03-01), pages 314-324, XP000910689 ISSN: 0014-2956 the whole document	1-27
A	MAEDA K ET AL: "FURTHER CHARACTERIZATION AND AMINO-ACID SEQUENCE OF M-TYPE THIOREDOXINS FROM SPINACH CHLOROPLASTS" EUROPEAN JOURNAL OF BIOCHEMISTRY 1986, vol. 154, no. 1, 1986, pages 197-204, XP000910501 ISSN: 0014-2956 the whole document	1-27
P,X	SHOEMAKER, R., ET AL. : "public soybean EST project" EMBL SEQUENCE DATA LIBRARY, 18 June 1999 (1999-06-18), XP002138132 heidelberg, germany accession no. AI736736	1-4,26

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 99/20420

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9603505 A	08-02-1996	FR 2723097 A	02-02-1996
		AU 3082795 A	22-02-1996
		EP 0802977 A	29-10-1997
WO 9519443 A	20-07-1995	US 5614395 A	25-03-1997
		AU 1249295 A	01-08-1995
		EP 0733117 A	25-09-1996
		US 5654414 A	05-08-1997
		US 5689044 A	18-11-1997
		US 5650505 A	22-07-1997
		US 5804693 A	08-09-1998
		US 5777200 A	07-07-1998
		US 5880328 A	09-03-1999
		US 5856154 A	05-01-1999
		US 5767369 A	16-06-1998
		US 5847258 A	08-12-1998
		US 5942662 A	24-08-1999
WO 9308274 A	29-04-1993	AT 175312 T	15-01-1999
		AU 677771 B	08-05-1997
		AU 2861792 A	21-05-1993
		CA 2121137 A	29-04-1993
		CZ 9400832 A	16-08-1995
		DE 69228130 D	18-02-1999
		EP 0672127 A	20-09-1995
		EP 0863154 A	09-09-1998
		HU 69780 A	28-09-1995
		JP 7502887 T	30-03-1995
		NZ 244695 A	26-03-1996
		SK 41894 A	11-07-1995
		US 5792506 A	11-08-1998
		US 5952034 A	14-09-1999
		ZA 9207831 A	27-04-1993
WO 9612799 A	02-05-1996	US 5792506 A	11-08-1998
		AU 3956395 A	15-05-1996
		EP 0784676 A	23-07-1997
		JP 10510244 T	06-10-1998
		US 5952034 A	14-09-1999

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

15

Applicant's or agent's file reference	FOR FURTHER ACTION	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
BB1246		
International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/US99/20420	07/09/1999	08/09/1998
International Patent Classification (IPC) or national classification and IPC C12N15/29		
Applicant E. I. DU PONT DE NEMOURS AND COMPANY et al.		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 7 sheets, including this cover sheet.

   ☐ This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

   These annexes consist of a total of sheets.

3. This report contains indications relating to the following items:

   - I ☒ Basis of the report
   - II ☐ Priority
   - III ☐ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
   - IV ☒ Lack of unity of invention
   - V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations suporting such statement
   - VI ☐ Certain documents cited
   - VII ☒ Certain defects in the international application
   - VIII ☒ Certain observations on the international application

Date of submission of the demand	Date of completion of this report
22/03/2000	05.12.2000
Name and mailing address of the international preliminary examining authority: European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465	Authorized officer SCHEFFZYK, I Telephone No. +49 89 2399 8602

Form PCT/IPEA/409 (cover sheet) (January 1994)

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/US99/20420

I. Basis of the report

1. This report has been drawn on the basis of *(substitute sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments (Rules 70.16 and 70.17).)*:
Description, pages:

1-22 as originally filed

Claims, No.:

1-27 as originally filed

2. With regard to the **language**, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.

These elements were available or furnished to this Authority in the following language: , which is:

- ☐ the language of a translation furnished for the purposes of the international search (under Rule 23.1(b)).
- ☐ the language of publication of the international application (under Rule 48.3(b)).
- ☐ the language of a translation furnished for the purposes of international preliminary examination (under Rule 55.2 and/or 55.3).

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:

- ☐ contained in the international application in written form.
- ☐ filed together with the international application in computer readable form.
- ☐ furnished subsequently to this Authority in written form.
- ☐ furnished subsequently to this Authority in computer readable form.
- ☐ The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.
- ☐ The statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished.

4. The amendments have resulted in the cancellation of:

- ☐ the description, pages:
- ☐ the claims, Nos.:
- ☐ the drawings, sheets:

5. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

(Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report.)

6. Additional observations, if necessary:

IV. Lack of unity of invention

1. In response to the invitation to restrict or pay additional fees the applicant has:

☐ restricted the claims.

☐ paid additional fees.

☐ paid additional fees under protest.

☒ neither restricted nor paid additional fees.

2. ☐ This Authority found that the requirement of unity of invention is not complied and chose, according to Rule 68.1, not to invite the applicant to restrict or pay additional fees.

3. This Authority considers that the requirement of unity of invention in accordance with Rules 13.1, 13.2 and 13.3 is

☐ complied with.

☒ not complied with for the following reasons:
see separate sheet

4. Consequently, the following parts of the international application were the subject of international preliminary examination in establishing this report:

☐ all parts.

☒ the parts relating to claims Nos. Claims 1-22 partly (relating to SEQ.ID.NOS. 1 and 2) and claim 23 complete.

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Yes:	Claims 1-14,16-20,23
	No:	Claims 15,21,22
Inventive step (IS)	Yes:	Claims 1-14,16-20,23
	No:	Claims 15,21,22
Industrial applicability (IA)	Yes:	Claims 1-23
	No:	Claims

2. Citations and explanations
see separate sheet

VII. Certain defects in the international application

The following defects in the form or contents of the international application have been noted:
see separate sheet

VIII. Certain observations on the international application

The following observations on the clarity of the claims, description, and drawings or on the question whether the claims are fully supported by the description, are made:
see separate sheet

SECTION IV----------

The common inventive concept underlying present application can be seen in the provision of polynucleotides isolated from plants encoding a thioredoxin polypeptide. However, considering that plant-specific nucleotide sequences that code for thioredoxins have already been disclosed in the prior art, (see for instance PNAS, 1995, 92, 5620-5624, Rivera-Madrid R. et al. or PMB, 1991, 17, 143-147, Marty I. and Mayer Y. said common inventive concept no longer exists. Moreover, considering the essential differences in the origin and primary structure of presently claimed thioredoxin sequences and considering also that there is no other technical feature common to the claimed subject-matter the IPEA is of the opinion that present application does not relate to one but to four separate inventions not linked by a common inventive concept as required by Rule 13.1-13.3 PCT, namely:

invention 1: Nucleotide sequence encoding a Momordica-specific thioredoxin polypeptide (SEQ.ID.NOS. 1 and 2), claim 23 complete and claims 1-22 partial;

invention 2: Nucleotide sequence encoding a Catalpa-specific thioredoxin polypeptide (SEQ.ID.NOS. 3 and 4), claim 24 complete and claims 1-22 partial

invention 3: Nucleotide sequence encoding a Soybean-specific thioredoxin polypeptide (SEQ.ID.NOS.: 5-8), claims 25 and 26 complete and claims 1-22 partial) and

invention 4: Nucleotide sequence encoding a Veronica-specific thioredoxin polypeptide (SEQ. ID.NOS. 9 and 10), claim 27 complete and claims 1-22 partial.

SECTION V----------

This IPER only relates to subject-matter directed to SEQ:ID.NO. 1 and 2 (see

section IV).

A polypeptide having the sequence shown in SEQ.ID.NO. 2 is neither taught nor suggested in the available prior art. Thus, claim 23 and claims 1-14 and 16-20 in so far as they relate to SEQ.I.NO. 2 or SEQ.ID.NO. 1 which shows the corresponding nucleotide sequence can be considered to be novel and inventive in the sense of Art. 33(2)(3) PCT.

Due to the expression "at least one of 30 contiguous nucleotides derived from " the method defined in claim 15 is not clearly delimitated from methods taught in the prior art (see e.g. WO-A-96/03505).

In addition, due to the unclear meaning of the term "substantial portion" used in claims 21 and 22 novelty of these claims vis-à-vis readily available methods of obtaining a nucleic acid fragment encoding a thioredoxin gene (see e.g. WO- A 96/03505) cannot be acknowledged. This is particularly true since the method defined in these claims covers the use of a sequence containing one nucleotide derived from the sequence shown in SEQ.ID.NO. 1 ("at least one of 40 or 30 contiguous nucleotides"). Thus, claims 15, 21 and 22 do not comply with the requirements of Art. 33(2)(3) PCT.

SECTION VII----------

The serial numbers should be replaced by the corresponding publication numbers.

SECTION VIII----------

1). In the absence of a second polypeptide the term "first" written in claims 1 and 23 appears to be superfluous.

2). The expression "substantial portion" used in claims 21 and 22 is relative and thus open to interpretation. Correspondingly, said expression renders the scope of these claims unclear (Art. 6 PCT).

3). The term "derived from" written in claims 15, 21 and 22 renders the scope of said claims unclear.

PCT 'ORLD INTELLECTUAL PROPERTY ORGANIZA
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/29, 15/53, 15/82, C12Q 1/68, C07K 14/415, C12N 9/02

A3

(11) International Publication Number: **WO 00/14239**

(43) International Publication Date: 16 March 2000 (16.03.00)

(21) International Application Number: PCT/US99/20420

(22) International Filing Date: 7 September 1999 (07.09.99)

(30) Priority Data:
60/099,501 8 September 1998 (08.09.98) US

(71) Applicant *(for all designated States except US)*: E.I. DU PONT DE NEMOURS AND COMPANY [US/US]; 1007 Market Street, Wilmington, DE 19898 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: ALLEN, Stephen, M. [US/US]; 2225 Rosewood Drive, Wilmington, DE 19810 (US). THORPE, Catherine, J. [GB/GB]; 20 The Beeches, Woodhead Drive, Cambridge CB4 1FY (GB). LU, Albert, L. [CA/US]; 14 Charles Pointe, Newark, DE 19702 (US).

(74) Agent: MAJARIAN, William, R.; E.I. du Pont de Nemours and Company, Legal Patent Records Center, 1007 Market Street, Wilmington, DE 19898 (US).

(81) Designated States: AE, AL, AU, BA, BB, BG, BR, CA, CN, CR, CU, CZ, EE, GD, GE, HR, HU, ID, IL, IN, IS, JP, KP, KR, LC, LK, LR, LT, LV, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, SL, TR, TT, UA, US, UZ, VN, YU, ZA, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(88) Date of publication of the international search report:
8 September 2000 (08.09.00)

(54) Title: THIOREDOXIN H HOMOLOGS

(57) Abstract

This invention relates to an isolated nucleic acid fragment encoding a thioredoxin protein. The invention also relates to the construction of a chimeric gene encoding all or a portion of the thioredoxin protein, in sense or antisense orientation, wherein expression of the chemeric gene results in production of altered levels of the thioredoxin protein in a transformed host cell.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZ
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/29, 15/53, 15/82, C12Q 1/68, C07K 14/415, C12N 9/02

A2

(11) International Publication Number: **WO 00/14239**

(43) International Publication Date: 16 March 2000 (16.03.00)

(21) International Application Number: PCT/US99/20420

(22) International Filing Date: 7 September 1999 (07.09.99)

(30) Priority Data:
60/099,501 8 September 1998 (08.09.98) US

(71) Applicant *(for all designated States except US)*: E.I. DU PONT DE NEMOURS AND COMPANY [US/US]; 1007 Market Street, Wilmington, DE 19898 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: ALLEN, Stephen, M. [US/US]; 2225 Rosewood Drive, Wilmington, DE 19810 (US). THORPE, Catherine, J. [GB/GB]; 20 The Beeches, Woodhead Drive, Cambridge CB4 1FY (GB). LU, Albert, L. [CA/US]; 14 Charles Pointe, Newark, DE 19702 (US).

(74) Agent: MAJARIAN, William, R.; E.I. du Pont de Nemours and Company, Legal Patent Records Center, 1007 Market Street, Wilmington, DE 19898 (US).

(81) Designated States: AE, AL, AU, BA, BB, BG, BR, CA, CN, CR, CU, CZ, EE, GD, GE, HR, HU, ID, IL, IN, IS, JP, KP, KR, LC, LK, LR, LT, LV, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, SL, TR, TT, UA, US, UZ, VN, YU, ZA, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: THIOREDOXIN H HOMOLOGS

(57) Abstract

This invention relates to an isolated nucleic acid fragment encoding a thioredoxin protein. The invention also relates to the construction of a chimeric gene encoding all or a portion of the thioredoxin protein, in sense or antisense orientation, wherein expression of the chemeric gene results in production of altered levels of the thioredoxin protein in a transformed host cell.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

TITLE

THIOREDOXIN H HOMOLOGS

This application claims the benefit of U.S. Provisional Application No. 60/099,501, filed September 8, 1998.

FIELD OF THE INVENTION

This invention is in the field of plant molecular biology. More specifically, this invention pertains to nucleic acid fragments encoding thioredoxin proteins in plants and seeds.

BACKGROUND OF THE INVENTION

Thioredoxin H is a cytosolic member of the thioredoxin family of proteins. These small proteins (typical mass of 12 kD) have been shown to play a central role in the activation of proteins by influencing the redox status of sulfhydryl groups on target proteins (Holmgren et al. (1985) *Annu. Rev. Biochem 54*:237-271; and Holmgren et al. (1995) *Structure 3(3)*:239-243). Two other thioredoxin classes, F and M, are located in plastids and have been shown to be involved in redox mediated activation/inactivation of various photosynthetic enzymes during light/dark transitions. The cytosolic (H) form of thioredoxin has been shown to be involved in disassembly of seed storage proteins during germination and in the bread making process. In the former case storage proteins are held together in clusters by S-S bonds. On germination thioredoxin H reduces the S-S bonds and the subunits dissociate, facilitating attack by proteases. During bread making the same processes occur. Reduction of the S-S bonds causes the protein complexes to disassemble allowing them to be distributed through out the dough during mixing. During kneading the S-H bonds become oxidized and start to reassociate in a random manner, the cross linked matrix formed by this process entraps CO_2 formed during yeast fermentation and is responsible for the raising process. Addition of thioredoxin H to poor quality flours improves their quality for the production of bread.

Thioredoxin H has also been shown to inactivate snake and bee venom toxins and has been shown to reduce the allergenicity of cereal proteins. In the later, the process is presumably the same as described above; by reducing the S-S bonds holding the storage protein clusters together they are more susceptible to denaturation and proteolysis in the gut. Thioredoxin H may also be overexpressed in transformed corn kernels and other cereal crops. The wet milling industry, which is primarily focused on starch extraction, steeps corn in liquors of sodium metabisulphite or SO_2. Although this has many secondary effects (e.g., suppression of microbial activity), the primary function is to cause a dissociation of the storage proteins which leads to more efficient starch extraction. Small increases in extractable starch translate into significant increases in the profit margins for the wet millers. By overexpressing thioredoxin H in maize kernels and other cereals it may possible to improve starch recoveries, reduce steep times, and reduce or eliminate the use of sulfur

compounds in the steeping process. Overexpression of thioredoxin H in maize kernels and other cereals may have the added advantage of reducing the allergenicity of any transgenic protein engineered into cereal crops with high sulfhydryl content.

SUMMARY OF THE INVENTION

The present invention relates to isolated polynucleotides comprising a nucleotide sequence encoding a first polypeptide of at least 100 amino acids that has at least 80% identity based on the Clustal Method of alignment when compared to a polypeptide selected from the group consisting of a *Momordica charantia* thioredoxin polypeptide of SEQ ID NO:2, a *Catakoa speciosa* thioredoxin polypeptide of SEQ ID NO:4, a soybean thioredoxin polypeptide of SEQ ID NO:6, a soybean thioredoxin polypeptide of SEQ ID NO:8, and a *Vernonia* thioredoxin polypeptide of SEQ ID NO:10. The present invention also relates to an isolated polynucleotide comprising the complement of the nucleotide sequences described above. It is preferred that the isolated polynucleotides of the claimed invention consists of regions of the isolated polynucleotide selected from the group SEQ ID NO:1, 3, 5, 7 and 9 that codes for the polypeptide selected from the group consisting of SEQ ID NO:2, 4, 6, 8 and 10. The present invention also relates to an isolated polynucleotide comprising a nucleotide sequences of at least one of 40 (preferably 30) contiguous nucleotides derived from a nucleotide sequence selected from the group consisting of SEQ ID NO:1, 3, 5, 7, 9 and the complement of such nucleotide sequences.

The present invention relates to a chimeric gene comprising an isolated polynucleotide of the present invention operably linked to suitable regulatory sequences.

The present invention relates to an isolated host cell comprising a chimeric gene of the present invention or an isolated polynucleotide of the present invention. The host cell may be eucaryotic, such as a yeast or a plant cell, or procaryotic, such as a bacterial cell. The present invention also relates to a virus, preferably a baculovirus, comprising an isolated polynucleotide of the present invention or a chimeric gene of the present invention.

The present invention relates to a process for producing an isolated host cell comprising a chimeric gene or isolated polynucleotide of the present invention, the process comprising either transforming or transfecting an isolated compatible host cell with a chimeric gene or isolated polynucleotide of the present invention.

The present invention relates to a thioredoxin polypeptide of at least 100 amino acids comprising at least 80% homology based on the Clustal method of alignment compared to a polypeptide selected from the group consisting of SEQ ID NO:2, 4, 6, 8 and 10.

The present invention relates to a method of selecting an isolated polynucleotide that affects the level of expression of a thioredoxin polypeptide in a plant cell, the method comprising the steps of:

constructing an isolated polynucleotide or chimeric gene of the present invention;

introducing the isolated polynucleotide into a plant cell;

measuring the level of thioredoxin polypeptide in the plant cell containing the polynucleotide; and

comparing the level of thioredoxin polypeptide in the plant cell containing the isolated polynucleotide with the level of thioredoxin polypeptide in a plant cell that does not contain the isolated polynucleotide.

The present invention relates to a method of obtaining a nucleic acid fragment encoding a substantial portion of a thioredoxin gene, preferably a plant thioredoxin gene, comprising the steps of: synthesizing an oligonucleotide primer comprising a nucleotide sequence of at least one of 40 (preferably 30) contiguous nucleotides derived from a nucleotide sequence selected from the group consisting of SEQ ID NO:1, 3, 5, 7, 9 and the complement of such nucleotide sequences; and amplifying a nucleic acid fragment (preferably a cDNA inserted in a cloning vector) using the oligonucleotide primer. The amplified nucleic acid fragment preferably will encode a portion of a thioredoxin polypeptide amino acid sequence.

The present invention also relates to a method of obtaining a nucleic acid fragment encoding all or a subsantial portion of the amino acid sequence encoding a thioredoxin protein comprising the steps of: probing a cDNA or genomic library with an isolated polynucleotide of the present invention; identifying a DNA clone that hybridizes with an isolated polynucleotide of the present invention; isolating the identified DNA clone; and sequencing the cDNA or genomic fragment that comprises the isolated DNA clone.

BRIEF DESCRIPTION OF THE SEQUENCE DESCRIPTIONS

The invention can be more fully understood from the following detailed description and the accompanying Sequence Listing which form a part of this application.

Table 1 lists the polypeptides that are described herein, the designation of the cDNA clones that comprise the nucleic acid fragments encoding polypeptides representing all or a substantial portion of these polypeptides, and the corresponding identifier (SEQ ID NO:) as used in the attached Sequence Listing. The sequence descriptions and Sequence Listing attached hereto comply with the rules governing nucleotide and/or amino acid sequence disclosures in patent applications as set forth in 37 C.F.R. §1.821-1.825.

TABLE 1

Thioredoxin Proteins

Protein	Clone Designation	SEQ ID NO: (Nucleotide)	(Amino Acid)
Thioredoxin H	fds.pk0001.e9	1	2
Thioredoxin H	ncs.pk0010.e3	3	4
Thioredoxin H	sah1c.pk001.l17	5	6
Thioredoxin H	sfl1.pk0029.e2	7	8
Thioredoxin H	vs1n.pk0012.f3	9	10

The Sequence Listing contains the one letter code for nucleotide sequence characters and the three letter codes for amino acids which is in conformity with the IUPAC-IUBMB standards described in *Nucleic Acids Res. 13*:3021-3030 (1985) and in the *Biochemical J. 219 (No. 2)*:345-373 (1984) which are herein incorporated by reference. The symbols and format used for nucleotide and amino acid sequence data comply with the rules set forth in 37 C.F.R. §1.822.

DETAILED DESCRIPTION OF THE INVENTION

In the context of this disclosure, a number of terms shall be utilized. As used herein, a "polynucleotide" is a nucleotide sequence such as a nucleic acid fragment. A polynucleotide may be a polymer of RNA or DNA that is single- or double-stranded, that optionally contains synthetic, non-natural or altered nucleotide bases. A polynucleotide in the form of a polymer of DNA may be comprised of one or more segments of cDNA, genomic DNA, or synthetic DNA.

As used herein, "substantially similar" refers to nucleic acid fragments wherein changes in one or more nucleotide bases results in substitution of one or more amino acids, but do not affect the functional properties of the polypeptide encoded by the nucleotide sequence. "Substantially similar" also refers to nucleic acid fragments wherein changes in one or more nucleotide bases does not affect the ability of the nucleic acid fragment to mediate alteration of gene expression by gene silencing through for example antisense or co-suppression technology. "Substantially similar" also refers to modifications of the nucleic acid fragments of the instant invention such as deletion or insertion of one or more nucleotides that do not substantially affect the functional properties of the resulting transcript vis-à-vis the ability to mediate gene silencing or alteration of the functional properties of the resulting protein molecule. It is therefore understood that the invention encompasses more than the specific exemplary nucleotide or amino acid sequences and includes functional equivalents thereof.

For example, it is well known in the art that antisense suppression and co-suppression of gene expression may be accomplished using nucleic acid fragments representing less than

the entire coding region of a gene, and by nucleic acid fragments that do not share 100% sequence identity with the gene to be suppressed. Moreover, alterations in a nucleic acid fragment which result in the production of a chemically equivalent amino acid at a given site, but do not effect the functional properties of the encoded polypeptide, are well known in the art. Thus, a codon for the amino acid alanine, a hydrophobic amino acid, may be substituted by a codon encoding another less hydrophobic residue, such as glycine, or a more hydrophobic residue, such as valine, leucine, or isoleucine. Similarly, changes which result in substitution of one negatively charged residue for another, such as aspartic acid for glutamic acid, or one positively charged residue for another, such as lysine for arginine, can also be expected to produce a functionally equivalent product. Nucleotide changes which result in alteration of the N-terminal and C-terminal portions of the polypeptide molecule would also not be expected to alter the activity of the polypeptide. Each of the proposed modifications is well within the routine skill in the art, as is determination of retention of biological activity of the encoded products. Consequently, a polynucleotide comprising a nucleotide sequence of at least one of 40 (preferably 30) contiguous nucleotides derived from a nucleotide sequence selected from the group consisting of SEQ ID NO:1, 3, 5, 7, 9 and the complement of such nucleotide sequences may be used in methods of selecting an isolated polynucleotide that affects the expression of a thioredoxin polypeptide in a plant cell.

Moreover, substantially similar nucleic acid fragments may also be characterized by their ability to hybridize. Estimates of such homology are provided by either DNA-DNA or DNA-RNA hybridization under conditions of stringency as is well understood by those skilled in the art (Hames and Higgins, Eds. (1985) Nucleic Acid Hybridisation, IRL Press, Oxford, U.K.). Stringency conditions can be adjusted to screen for moderately similar fragments, such as homologous sequences from distantly related organisms, to highly similar fragments, such as genes that duplicate functional enzymes from closely related organisms. Post-hybridization washes determine stringency conditions. One set of preferred conditions uses a series of washes starting with 6X SSC, 0.5% SDS at room temperature for 15 min, then repeated with 2X SSC, 0.5% SDS at 45°C for 30 min, and then repeated twice with 0.2X SSC, 0.5% SDS at 50°C for 30 min. A more preferred set of stringent conditions uses higher temperatures in which the washes are identical to those above except for the temperature of the final two 30 min washes in 0.2X SSC, 0.5% SDS was increased to 60°C. Another preferred set of highly stringent conditions uses two final washes in 0.1X SSC, 0.1% SDS at 65°C.

Substantially similar nucleic acid fragments of the instant invention may also be characterized by the percent identity of the amino acid sequences that they encode to the amino acid sequences disclosed herein, as determined by algorithms commonly employed by those skilled in this art. Suitable nucleic acid fragments (polynucleotides) encode amino

acid sequences that are 80% identical to the amino acid sequences reported herein. Preferred nucleic acid fragment encode amino acid sequences that are 85% identical to the amino acid sequences reported herein. More preferred nucleic acid fragments encode amino acid sequences that are 90% identical to the amino acid sequences reported herein. Most preferred are nucleic acid fragments that encode amino acid sequences that are 95% identical to the amino acid sequences reported herein. Sequence alignments and percent identity calculations were performed using the Megalign program of the LASERGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Multiple alignment of the sequences was performed using the Clustal method of alignment (Higgins and Sharp (1989) *CABIOS. 5*:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default parameters for pairwise alignments using the Clustal method were KTUPLE 1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5.

A "substantial portion" of an amino acid or nucleotide sequence comprises an amino acid or a nucleotide sequence that is sufficient to afford putative identification of the protein or gene that the amino acid or nucleotide sequence comprises. Amino acid and nucleotide sequences can be evaluated either manually by one skilled in the art, or by using computer-based sequence comparison and identification tools that employ algorithms such as BLAST (Basic Local Alignment Search Tool; Altschul et al. (1993) *J. Mol. Biol. 215*:403-410; see also www.ncbi.nlm.nih.gov/BLAST/). In general, a sequence of ten or more contiguous amino acids or thirty or more contiguous nucleotides is necessary in order to putatively identify a polypeptide or nucleic acid sequence as homologous to a known protein or gene. Moreover, with respect to nucleotide sequences, gene-specific oligonucleotide probes comprising 30 or more contiguous nucleotides may be used in sequence-dependent methods of gene identification (e.g., Southern hybridization) and isolation (e.g., *in situ* hybridization of bacterial colonies or bacteriophage plaques). In addition, short oligonucleotides of 12 or more nucleotides may be used as amplification primers in PCR in order to obtain a particular nucleic acid fragment comprising the primers. Accordingly, a "substantial portion" of a nucleotide sequence comprises a nucleotide sequence that will afford specific identification and/or isolation of a nucleic acid fragment comprising the sequence. The instant specification teaches amino acid and nucleotide sequences encoding polypeptides that comprise one or more particular plant proteins. The skilled artisan, having the benefit of the sequences as reported herein, may now use all or a substantial portion of the disclosed sequences for purposes known to those skilled in this art. Accordingly, the instant invention comprises the complete sequences as reported in the accompanying Sequence Listing, as well as substantial portions of those sequences as defined above.

"Codon degeneracy" refers to divergence in the genetic code permitting variation of the nucleotide sequence without effecting the amino acid sequence of an encoded polypeptide. Accordingly, the instant invention relates to any nucleic acid fragment

comprising a nucleotide sequence that encodes all or a substantial portion of the amino acid sequences set forth herein. The skilled artisan is well aware of the "codon-bias" exhibited by a specific host cell in usage of nucleotide codons to specify a given amino acid. Therefore, when synthesizing a nucleic acid fragment for improved expression in a host cell, it is desirable to design the nucleic acid fragment such that its frequency of codon usage approaches the frequency of preferred codon usage of the host cell.

"Synthetic nucleic acid fragments" can be assembled from oligonucleotide building blocks that are chemically synthesized using procedures known to those skilled in the art. These building blocks are ligated and annealed to form larger nucleic acid fragments which may then be enzymatically assembled to construct the entire desired nucleic acid fragment. "Chemically synthesized", as related to nucleic acid fragment, means that the component nucleotides were assembled *in vitro*. Manual chemical synthesis of nucleic acid fragments may be accomplished using well established procedures, or automated chemical synthesis can be performed using one of a number of commercially available machines. Accordingly, the nucleic acid fragments can be tailored for optimal gene expression based on optimization of nucleotide sequence to reflect the codon bias of the host cell. The skilled artisan appreciates the likelihood of successful gene expression if codon usage is biased towards those codons favored by the host. Determination of preferred codons can be based on a survey of genes derived from the host cell where sequence information is available.

"Gene" refers to a nucleic acid fragment that expresses a specific protein, including regulatory sequences preceding (5' non-coding sequences) and following (3' non-coding sequences) the coding sequence. "Native gene" refers to a gene as found in nature with its own regulatory sequences. "Chimeric gene" refers any gene that is not a native gene, comprising regulatory and coding sequences that are not found together in nature. Accordingly, a chimeric gene may comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived from the same source, but arranged in a manner different than that found in nature. "Endogenous gene" refers to a native gene in its natural location in the genome of an organism. A "foreign" gene refers to a gene not normally found in the host organism, but that is introduced into the host organism by gene transfer. Foreign genes can comprise native genes inserted into a non-native organism, or chimeric genes. A "transgene" is a gene that has been introduced into the genome by a transformation procedure.

"Coding sequence" refers to a nucleotide sequence that codes for a specific amino acid sequence. "Regulatory sequences" refer to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences may include promoters, translation leader sequences, introns, and polyadenylation recognition sequences.

"Promoter" refers to a nucleotide sequence capable of controlling the expression of a coding sequence or functional RNA. In general, a coding sequence is located 3' to a promoter sequence. The promoter sequence consists of proximal and more distal upstream elements, the latter elements often referred to as enhancers. Accordingly, an "enhancer" is a nucleotide sequence which can stimulate promoter activity and may be an innate element of the promoter or a heterologous element inserted to enhance the level or tissue-specificity of a promoter. Promoters may be derived in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or even comprise synthetic nucleotide segments. It is understood by those skilled in the art that different promoters may direct the expression of a gene in different tissues or cell types, or at different stages of development, or in response to different environmental conditions. Promoters which cause a nucleic acid fragment to be expressed in most cell types at most times are commonly referred to as "constitutive promoters". New promoters of various types useful in plant cells are constantly being discovered; numerous examples may be found in the compilation by Okamuro and Goldberg (1989) *Biochemistry of Plants 15*:1-82. It is further recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, nucleic acid fragments of different lengths may have identical promoter activity.

The "translation leader sequence" refers to a nucleotide sequence located between the promoter sequence of a gene and the coding sequence. The translation leader sequence is present in the fully processed mRNA upstream of the translation start sequence. The translation leader sequence may affect processing of the primary transcript to mRNA, mRNA stability or translation efficiency. Examples of translation leader sequences have been described (Turner and Foster (1995) *Mol. Biotechnol. 3*:225-236).

The "3' non-coding sequences" refer to nucleotide sequences located downstream of a coding sequence and include polyadenylation recognition sequences and other sequences encoding regulatory signals capable of affecting mRNA processing or gene expression. The polyadenylation signal is usually characterized by affecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor. The use of different 3' non-coding sequences is exemplified by Ingelbrecht et al. (1989) *Plant Cell 1*:671-680.

"RNA transcript" refers to the product resulting from RNA polymerase-catalyzed transcription of a DNA sequence. When the RNA transcript is a perfect complementary copy of the DNA sequence, it is referred to as the primary transcript or it may be a RNA sequence derived from posttranscriptional processing of the primary transcript and is referred to as the mature RNA. "Messenger RNA (mRNA)" refers to the RNA that is without introns and that can be translated into polypeptide by the cell. "cDNA" refers to a double-stranded DNA that is complementary to and derived from mRNA. "Sense" RNA refers to an RNA transcript that includes the mRNA and so can be translated into a

polypeptide by the cell. "Antisense RNA" refers to an RNA transcript that is complementary to all or part of a target primary transcript or mRNA and that blocks the expression of a target gene (see U.S. Patent No. 5,107,065, incorporated herein by reference). The complementarity of an antisense RNA may be with any part of the specific nucleotide sequence, i.e., at the 5' non-coding sequence, 3' non-coding sequence, introns, or the coding sequence. "Functional RNA" refers to sense RNA, antisense RNA, ribozyme RNA, or other RNA that may not be translated but yet has an effect on cellular processes.

The term "operably linked" refers to the association of two or more nucleic acid fragments on a single nucleic acid fragment so that the function of one is affected by the other. For example, a promoter is operably linked with a coding sequence when it is capable of affecting the expression of that coding sequence (i.e., that the coding sequence is under the transcriptional control of the promoter). Coding sequences can be operably linked to regulatory sequences in sense or antisense orientation.

The term "expression", as used herein, refers to the transcription and stable accumulation of sense (mRNA) or antisense RNA derived from the nucleic acid fragment of the invention. Expression may also refer to translation of mRNA into a polypeptide. "Antisense inhibition" refers to the production of antisense RNA transcripts capable of suppressing the expression of the target protein. "Overexpression" refers to the production of a gene product in transgenic organisms that exceeds levels of production in normal or non-transformed organisms. "Co-suppression" refers to the production of sense RNA transcripts capable of suppressing the expression of identical or substantially similar foreign or endogenous genes (U.S. Patent No. 5,231,020, incorporated herein by reference).

"Altered levels" refers to the production of gene product(s) in transgenic organisms in amounts or proportions that differ from that of normal or non-transformed organisms.

"Mature" protein refers to a post-translationally processed polypeptide; i.e., one from which any pre- or propeptides present in the primary translation product have been removed. "Precursor" protein refers to the primary product of translation of mRNA; i.e., with pre- and propeptides still present. Pre- and propeptides may be but are not limited to intracellular localization signals.

A "chloroplast transit peptide" is an amino acid sequence which is translated in conjunction with a protein and directs the protein to the chloroplast or other plastid types present in the cell in which the protein is made. "Chloroplast transit sequence" refers to a nucleotide sequence that encodes a chloroplast transit peptide. A "signal peptide" is an amino acid sequence which is translated in conjunction with a protein and directs the protein to the secretory system (Chrispeels (1991) *Ann. Rev. Plant Phys. Plant Mol. Biol. 42*:21-53). If the protein is to be directed to a vacuole, a vacuolar targeting signal (*supra*) can further be added, or if to the endoplasmic reticulum, an endoplasmic reticulum retention signal (*supra*) may be added. If the protein is to be directed to the nucleus, any signal peptide present

should be removed and instead a nuclear localization signal included (Raikhel (1992) *Plant Phys. 100*:1627-1632).

"Transformation" refers to the transfer of a nucleic acid fragment into the genome of a host organism, resulting in genetically stable inheritance. Host organisms containing the transformed nucleic acid fragments are referred to as "transgenic" organisms. Examples of methods of plant transformation include *Agrobacterium*-mediated transformation (De Blaere et al. (1987) *Meth. Enzymol. 143*:277) and particle-accelerated or "gene gun" transformation technology (Klein et al. (1987) *Nature (London) 327*:70-73; U.S. Patent No. 4,945,050, incorporated herein by reference).

Standard recombinant DNA and molecular cloning techniques used herein are well known in the art and are described more fully in Sambrook et al. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 1989 (hereinafter "Maniatis").

Nucleic acid fragments encoding at least a portion of several thioredoxin proteins have been isolated and identified by comparison of random plant cDNA sequences to public databases containing nucleotide and protein sequences using the BLAST algorithms well known to those skilled in the art. The nucleic acid fragments of the instant invention may be used to isolate cDNAs and genes encoding homologous proteins from the same or other plant species. Isolation of homologous genes using sequence-dependent protocols is well known in the art. Examples of sequence-dependent protocols include, but are not limited to, methods of nucleic acid hybridization, and methods of DNA and RNA amplification as exemplified by various uses of nucleic acid amplification technologies (e.g., polymerase chain reaction, ligase chain reaction).

For example, genes encoding other thioredoxin H proteins, either as cDNAs or genomic DNAs, could be isolated directly by using all or a portion of the instant nucleic acid fragments as DNA hybridization probes to screen libraries from any desired plant employing methodology well known to those skilled in the art. Specific oligonucleotide probes based upon the instant nucleic acid sequences can be designed and synthesized by methods known in the art (Maniatis). Moreover, the entire sequences can be used directly to synthesize DNA probes by methods known to the skilled artisan such as random primer DNA labeling, nick translation, or end-labeling techniques, or RNA probes using available *in vitro* transcription systems. In addition, specific primers can be designed and used to amplify a part or all of the instant sequences. The resulting amplification products can be labeled directly during amplification reactions or labeled after amplification reactions, and used as probes to isolate full length cDNA or genomic fragments under conditions of appropriate stringency.

In addition, two short segments of the instant nucleic acid fragments may be used in polymerase chain reaction protocols to amplify longer nucleic acid fragments encoding

homologous genes from DNA or RNA. The polymerase chain reaction may also be performed on a library of cloned nucleic acid fragments wherein the sequence of one primer is derived from the instant nucleic acid fragments, and the sequence of the other primer takes advantage of the presence of the polyadenylic acid tracts to the 3' end of the mRNA precursor encoding plant genes. Alternatively, the second primer sequence may be based upon sequences derived from the cloning vector. For example, the skilled artisan can follow the RACE protocol (Frohman et al. (1988) *Proc. Natl. Acad. Sci. USA 85*:8998-9002) to generate cDNAs by using PCR to amplify copies of the region between a single point in the transcript and the 3' or 5' end. Primers oriented in the 3' and 5' directions can be designed from the instant sequences. Using commercially available 3' RACE or 5' RACE systems (BRL), specific 3' or 5' cDNA fragments can be isolated (Ohara et al. (1989) *Proc. Natl. Acad. Sci. USA 86*:5673-5677; Loh et al. (1989) *Science 243*:217-220). Products generated by the 3' and 5' RACE procedures can be combined to generate full-length cDNAs (Frohman and Martin (1989) *Techniques 1*:165). Consequently, a polynucleotide comprising a nucleotide sequence of at least one of 40 (preferably 30) contiguous nucleotides derived from a nucleotide sequence selected from the group consisting of SEQ ID NO:1, 3, 5, 7, 9 and the complement of such nucleotide sequences may be used in such methods to obtain a nucleic acid fragment encoding a substantial portion of an amino acid sequence of a thioredoxin polypeptide.

Availability of the instant nucleotide and deduced amino acid sequences facilitates immunological screening of cDNA expression libraries. Synthetic peptides representing portions of the instant amino acid sequences may be synthesized. These peptides can be used to immunize animals to produce polyclonal or monoclonal antibodies with specificity for peptides or proteins comprising the amino acid sequences. These antibodies can be then be used to screen cDNA expression libraries to isolate full-length cDNA clones of interest (Lerner (1984) *Adv. Immunol. 36*:1-34; Maniatis).

The nucleic acid fragments of the instant invention may be used to create transgenic plants in which the disclosed polypeptides are present at higher or lower levels than normal or in cell types or developmental stages in which they are not normally found. This would have the effect of altering the level of thioredoxin activity in those cells.

Overexpression of the proteins of the instant invention may be accomplished by first constructing a chimeric gene in which the coding region is operably linked to a promoter capable of directing expression of a gene in the desired tissues at the desired stage of development. For reasons of convenience, the chimeric gene may comprise promoter sequences and translation leader sequences derived from the same genes. 3' Non-coding sequences encoding transcription termination signals may also be provided. The instant chimeric gene may also comprise one or more introns in order to facilitate gene expression.

Plasmid vectors comprising the instant chimeric gene can then be constructed. The choice of plasmid vector is dependent upon the method that will be used to transform host plants. The skilled artisan is well aware of the genetic elements that must be present on the plasmid vector in order to successfully transform, select and propagate host cells containing the chimeric gene. The skilled artisan will also recognize that different independent transformation events will result in different levels and patterns of expression (Jones et al. (1985) *EMBO J. 4*:2411-2418; De Almeida et al. (1989) *Mol. Gen. Genetics 218*:78-86), and thus that multiple events must be screened in order to obtain lines displaying the desired expression level and pattern. Such screening may be accomplished by Southern analysis of DNA, Northern analysis of mRNA expression, Western analysis of protein expression, or phenotypic analysis.

For some applications it may be useful to direct the instant polypeptides to different cellular compartments, or to facilitate its secretion from the cell. It is thus envisioned that the chimeric gene described above may be further supplemented by altering the coding sequence to encode the instant polypeptides with appropriate intracellular targeting sequences such as transit sequences (Keegstra (1989) *Cell 56*:247-253), signal sequences or sequences encoding endoplasmic reticulum localization (Chrispeels (1991) *Ann. Rev. Plant Phys. Plant Mol. Biol. 42*:21-53), or nuclear localization signals (Raikhel (1992) *Plant Phys.100*:1627-1632) added and/or with targeting sequences that are already present removed. While the references cited give examples of each of these, the list is not exhaustive and more targeting signals of utility may be discovered in the future.

It may also be desirable to reduce or eliminate expression of genes encoding the instant polypeptides in plants for some applications. In order to accomplish this, a chimeric gene designed for co-suppression of the instant polypeptide can be constructed by linking a gene or gene fragment encoding that polypeptide to plant promoter sequences. Alternatively, a chimeric gene designed to express antisense RNA for all or part of the instant nucleic acid fragment can be constructed by linking the gene or gene fragment in reverse orientation to plant promoter sequences. Either the co-suppression or antisense chimeric genes could be introduced into plants via transformation wherein expression of the corresponding endogenous genes are reduced or eliminated.

Molecular genetic solutions to the generation of plants with altered gene expression have a decided advantage over more traditional plant breeding approaches. Changes in plant phenotypes can be produced by specifically inhibiting expression of one or more genes by antisense inhibition or cosuppression (U.S. Patent Nos. 5,190,931, 5,107,065 and 5,283,323). An antisense or cosuppression construct would act as a dominant negative regulator of gene activity. While conventional mutations can yield negative regulation of gene activity these effects are most likely recessive. The dominant negative regulation available with a transgenic approach may be advantageous from a breeding perspective. In

addition, the ability to restrict the expression of specific phenotype to the reproductive tissues of the plant by the use of tissue specific promoters may confer agronomic advantages relative to conventional mutations which may have an effect in all tissues in which a mutant gene is ordinarily expressed.

The person skilled in the art will know that special considerations are associated with the use of antisense or cosuppression technologies in order to reduce expression of particular genes. For example, the proper level of expression of sense or antisense genes may require the use of different chimeric genes utilizing different regulatory elements known to the skilled artisan. Once transgenic plants are obtained by one of the methods described above, it will be necessary to screen individual transgenics for those that most effectively display the desired phenotype. Accordingly, the skilled artisan will develop methods for screening large numbers of transformants. The nature of these screens will generally be chosen on practical grounds, and is not an inherent part of the invention. For example, one can screen by looking for changes in gene expression by using antibodies specific for the protein encoded by the gene being suppressed, or one could establish assays that specifically measure enzyme activity. A preferred method will be one which allows large numbers of samples to be processed rapidly, since it will be expected that a large number of transformants will be negative for the desired phenotype.

The instant polypeptides (or portions thereof) may be produced in heterologous host cells, particularly in the cells of microbial hosts, and can be used to prepare antibodies to the these proteins by methods well known to those skilled in the art. The antibodies are useful for detecting the polypeptides of the instant invention *in situ* in cells or *in vitro* in cell extracts. Preferred heterologous host cells for production of the instant polypeptides are microbial hosts. Microbial expression systems and expression vectors containing regulatory sequences that direct high level expression of foreign proteins are well known to those skilled in the art. Any of these could be used to construct a chimeric gene for production of the instant polypeptides. This chimeric gene could then be introduced into appropriate microorganisms via transformation to provide high level expression of the encoded thioredoxin protein. An example of a vector for high level expression of the instant polypeptides in a bacterial host is provided (Example 6).

All or a substantial portion of the nucleic acid fragments of the instant invention may also be used as probes for genetically and physically mapping the genes that they are a part of, and as markers for traits linked to those genes. Such information may be useful in plant breeding in order to develop lines with desired phenotypes. For example, the instant nucleic acid fragments may be used as restriction fragment length polymorphism (RFLP) markers. Southern blots (Maniatis) of restriction-digested plant genomic DNA may be probed with the nucleic acid fragments of the instant invention. The resulting banding patterns may then be subjected to genetic analyses using computer programs such as MapMaker (Lander et al.

(1987) *Genomics 1*:174-181) in order to construct a genetic map. In addition, the nucleic acid fragments of the instant invention may be used to probe Southern blots containing restriction endonuclease-treated genomic DNAs of a set of individuals representing parent and progeny of a defined genetic cross. Segregation of the DNA polymorphisms is noted and used to calculate the position of the instant nucleic acid sequence in the genetic map previously obtained using this population (Botstein et al. (1980) *Am. J. Hum. Genet. 32*:314-331).

The production and use of plant gene-derived probes for use in genetic mapping is described in Bernatzky and Tanksley (1986) *Plant Mol. Biol. Reporter 4*:37-41. Numerous publications describe genetic mapping of specific cDNA clones using the methodology outlined above or variations thereof. For example, F2 intercross populations, backcross populations, randomly mated populations, near isogenic lines, and other sets of individuals may be used for mapping. Such methodologies are well known to those skilled in the art.

Nucleic acid probes derived from the instant nucleic acid sequences may also be used for physical mapping (i.e., placement of sequences on physical maps; *see* Hoheisel et al. In: *Nonmammalian Genomic Analysis: A Practical Guide*, Academic press 1996, pp. 319-346, and references cited therein).

In another embodiment, nucleic acid probes derived from the instant nucleic acid sequences may be used in direct fluorescence *in situ* hybridization (FISH) mapping (Trask (1991) *Trends Genet. 7*:149-154). Although current methods of FISH mapping favor use of large clones (several to several hundred KB; see Laan et al. (1995) *Genome Res. 5*:13-20), improvements in sensitivity may allow performance of FISH mapping using shorter probes.

A variety of nucleic acid amplification-based methods of genetic and physical mapping may be carried out using the instant nucleic acid sequences. Examples include allele-specific amplification (Kazazian (1989) *J. Lab. Clin. Med. 11*:95-96), polymorphism of PCR-amplified fragments (CAPS; Sheffield et al. (1993) *Genomics 16*:325-332), allele-specific ligation (Landegren et al. (1988) *Science 241*:1077-1080), nucleotide extension reactions (Sokolov (1990) *Nucleic Acid Res. 18*:3671), Radiation Hybrid Mapping (Walter et al. (1997) *Nat. Genet. 7*:22-28) and Happy Mapping (Dear and Cook (1989) *Nucleic Acid Res. 17*:6795-6807). For these methods, the sequence of a nucleic acid fragment is used to design and produce primer pairs for use in the amplification reaction or in primer extension reactions. The design of such primers is well known to those skilled in the art. In methods employing PCR-based genetic mapping, it may be necessary to identify DNA sequence differences between the parents of the mapping cross in the region corresponding to the instant nucleic acid sequence. This, however, is generally not necessary for mapping methods.

Loss of function mutant phenotypes may be identified for the instant cDNA clones either by targeted gene disruption protocols or by identifying specific mutants for these

genes contained in a maize population carrying mutations in all possible genes (Ballinger and Benzer (1989) *Proc. Natl. Acad. Sci USA 86*:9402-9406; Koes et al. (1995) *Proc. Natl. Acad. Sci USA 92*:8149-8153; Bensen et al. (1995) *Plant Cell 7*:75-84). The latter approach may be accomplished in two ways. First, short segments of the instant nucleic acid fragments may be used in polymerase chain reaction protocols in conjunction with a mutation tag sequence primer on DNAs prepared from a population of plants in which Mutator transposons or some other mutation-causing DNA element has been introduced (see Bensen, *supra*). The amplification of a specific DNA fragment with these primers indicates the insertion of the mutation tag element in or near the plant gene encoding the instant polypeptides. Alternatively, the instant nucleic acid fragment may be used as a hybridization probe against PCR amplification products generated from the mutation population using the mutation tag sequence primer in conjunction with an arbitrary genomic site primer, such as that for a restriction enzyme site-anchored synthetic adaptor. With either method, a plant containing a mutation in the endogenous gene encoding the instant polypeptides can be identified and obtained. This mutant plant can then be used to determine or confirm the natural function of the instant polypeptides disclosed herein.

EXAMPLES

The present invention is further defined in the following Examples, in which all parts and percentages are by weight and degrees are Celsius, unless otherwise stated. It should be understood that these Examples, while indicating preferred embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions.

EXAMPLE 1

Composition of cDNA Libraries; Isolation and Sequencing of cDNA Clones

cDNA libraries representing mRNAs from various barley, *Catalpa*, pear, soybean and *Vernonia* tissues were prepared. The characteristics of the libraries are described below.

TABLE 2
cDNA Libraries from Barley, *Catalpa*, Pear, Soybean and *Vernonia*

Library	Tissue	Clone
fds	*Momordica charantia* developing seed	fds.pk0001.e9
ncs	*Catalpa speciosa* developing Seed	ncs.pk0010.e3
sah1c	Soybean sprayed with Authority herbicide.	sah1c.pk001.l17
sfl1	Soybean immature flower	sfl1.pk0029.e2
vs1n	Vernonia Seed*	vs1n.pk0012.f3

*This library was normalized essentially as described in U.S. Patent No. 5,482,845, incorporated herein by reference.

cDNA libraries may be prepared by any one of many methods available. For example, the cDNAs may be introduced into plasmid vectors by first preparing the cDNA libraries in Uni-ZAP™ XR vectors according to the manufacturer's protocol (Stratagene Cloning Systems, La Jolla, CA). The Uni-ZAP™ XR libraries are converted into plasmid libraries according to the protocol provided by Stratagene. Upon conversion, cDNA inserts will be contained in the plasmid vector pBluescript. In addition, the cDNAs may be introduced directly into precut Bluescript II SK(+) vectors (Stratagene) using T4 DNA ligase (New England Biolabs), followed by transfection into DH10B cells according to the manufacturer's protocol (GIBCO BRL Products). Once the cDNA inserts are in plasmid vectors, plasmid DNAs are prepared from randomly picked bacterial colonies containing recombinant pBluescript plasmids, or the insert cDNA sequences are amplified via polymerase chain reaction using primers specific for vector sequences flanking the inserted cDNA sequences. Amplified insert DNAs or plasmid DNAs are sequenced in dye-primer sequencing reactions to generate partial cDNA sequences (expressed sequence tags or "ESTs"; see Adams et al., (1991) *Science 252*:1651-1656). The resulting ESTs are analyzed using a Perkin Elmer Model 377 fluorescent sequencer.

EXAMPLE 2

Identification of cDNA Clones

cDNA clones encoding thioredoxin proteins were identified by conducting BLAST (Basic Local Alignment Search Tool; Altschul et al. (1993) *J. Mol. Biol. 215*:403-410; see also www.ncbi.nlm.nih.gov/BLAST/) searches for similarity to sequences contained in the BLAST "nr" database (comprising all non-redundant GenBank CDS translations, sequences derived from the 3-dimensional structure Brookhaven Protein Data Bank, the last major release of the SWISS-PROT protein sequence database, EMBL, and DDBJ databases). The cDNA sequences obtained in Example 1 were analyzed for similarity to all publicly available DNA sequences contained in the "nr" database using the BLASTN algorithm provided by the National Center for Biotechnology Information (NCBI). The DNA sequences were translated in all reading frames and compared for similarity to all publicly available protein sequences contained in the "nr" database using the BLASTX algorithm (Gish and States (1993) *Nat. Genet. 3*:266-272) provided by the NCBI. For convenience, the P-value (probability) of observing a match of a cDNA sequence to a sequence contained in the searched databases merely by chance as calculated by BLAST are reported herein as "pLog" values, which represent the negative of the logarithm of the reported P-value. Accordingly, the greater the pLog value, the greater the likelihood that the cDNA sequence and the BLAST "hit" represent homologous proteins.

EXAMPLE 3

Characterization of cDNA Clones Encoding Thioredoxin H

The BLASTX search using the EST sequences from clones listed in Table 3 revealed similarity of the polypeptides encoded by the cDNAs to thioredoxin H from *Arabidopsis thaliana* (NCBI Identifier No. gi 267122), *Nicotiana tabacum* (NCBI Identifier No. gi 267124) and *Ricinus communis* (NCBI Identifier No. gi 1255954). Shown in Table 3 are the BLAST results for individual ESTs ("EST"), the sequences of the entire cDNA inserts comprising the indicated cDNA clones ("FIS"), or contigs assembled from two or more ESTs ("Contig"):

TABLE 3

BLAST Results for Sequences Encoding Polypeptides Homologous to *Arabidopsis thaliana, Nicotiana tabacum* and *Ricinus communis* Thioredoxin H

Clone	Status	BLAST pLog Score
fds.pk0001.e9	FIS	35.50 (gi 267122)
ncs.pk0010.e3	FIS	48.40 (gi 267124)
sah1c.pk001.l17	FIS	49.00 (gi 1255954)
sfl1.pk0029.e2	FIS	41.00 (gi 1255954)
vs1n.pk0012.f3	FIS	41.70 (gi 267124)

The data in Table 4 represents a calculation of the percent identity of the amino acid sequences set forth in SEQ ID NOs:2, 4, 6, 8 and 10 and the *Arabidopsis thaliana, Nicotiana tabacum* and *Ricinus communis* sequences (SEQ ID NOs:11, 12 and 13 respectively). The percent identity between the amino acid sequences set forth in SEQ ID NOs:2, 4, 6, 8 and 10 ranged from 49% to 80%.

TABLE 4

Percent Identity of Amino Acid Sequences Deduced From the Nucleotide Sequences of cDNA Clones Encoding Polypeptides Homologous to *Arabidopsis thaliana, Nicotiana tabacum* and *Ricinus communis* Thioredoxin H

SEQ ID NO.	Percent Identity to
2	62% (gi 267122)
4	75% (gi 267124)
6	75% (gi 1255954)
8	65% (gi 1255954)
10	69% (gi 267124)

Sequence alignments and percent identity calculations were performed using the Megalign program of the LASERGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Multiple alignment of the sequences was performed using the Clustal

method of alignment (Higgins and Sharp (1989) *CABIOS.* 5:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default parameters for pairwise alignments using the Clustal method were KTUPLE 1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5. Sequence alignments and BLAST scores and probabilities indicate that the nucleic acid fragments comprising the instant cDNA clones encode a substantial portion of a thioredoxin H. These sequences represent the first *Catalpa*, pear, soybean and *Vernonia* sequences encoding thioredoxin H.

EXAMPLE 4

Expression of Chimeric Genes in Monocot Cells

A chimeric gene comprising a cDNA encoding the instant polypeptides in sense orientation with respect to the maize 27 kD zein promoter that is located 5' to the cDNA fragment, and the 10 kD zein 3' end that is located 3' to the cDNA fragment, can be constructed. The cDNA fragment of this gene may be generated by polymerase chain reaction (PCR) of the cDNA clone using appropriate oligonucleotide primers. Cloning sites (NcoI or SmaI) can be incorporated into the oligonucleotides to provide proper orientation of the DNA fragment when inserted into the digested vector pML103 as described below. Amplification is then performed in a standard PCR. The amplified DNA is then digested with restriction enzymes NcoI and SmaI and fractionated on an agarose gel. The appropriate band can be isolated from the gel and combined with a 4.9 kb NcoI-SmaI fragment of the plasmid pML103. Plasmid pML103 has been deposited under the terms of the Budapest Treaty at ATCC (American Type Culture Collection, 10801 University Blvd., Manassas, VA 20110-2209), and bears accession number ATCC 97366. The DNA segment from pML103 contains a 1.05 kb SalI-NcoI promoter fragment of the maize 27 kD zein gene and a 0.96 kb SmaI-SalI fragment from the 3' end of the maize 10 kD zein gene in the vector pGem9Zf(+) (Promega). Vector and insert DNA can be ligated at 15°C overnight, essentially as described (Maniatis). The ligated DNA may then be used to transform *E. coli* XL1-Blue (Epicurian Coli XL-1 Blue™; Stratagene). Bacterial transformants can be screened by restriction enzyme digestion of plasmid DNA and limited nucleotide sequence analysis using the dideoxy chain termination method (Sequenase™ DNA Sequencing Kit; U.S. Biochemical). The resulting plasmid construct would comprise a chimeric gene encoding, in the 5' to 3' direction, the maize 27 kD zein promoter, a cDNA fragment encoding the instant polypeptides, and the 10 kD zein 3' region.

The chimeric gene described above can then be introduced into corn cells by the following procedure. Immature corn embryos can be dissected from developing caryopses derived from crosses of the inbred corn lines H99 and LH132. The embryos are isolated 10 to 11 days after pollination when they are 1.0 to 1.5 mm long. The embryos are then placed with the axis-side facing down and in contact with agarose-solidified N6 medium (Chu et al. (1975) *Sci. Sin. Peking* 18:659-668). The embryos are kept in the dark at 27°C. Friable

embryogenic callus consisting of undifferentiated masses of cells with somatic proembryoids and embryoids borne on suspensor structures proliferates from the scutellum of these immature embryos. The embryogenic callus isolated from the primary explant can be cultured on N6 medium and sub-cultured on this medium every 2 to 3 weeks.

The plasmid, p35S/Ac (obtained from Dr. Peter Eckes, Hoechst Ag, Frankfurt, Germany) may be used in transformation experiments in order to provide for a selectable marker. This plasmid contains the *Pat* gene (see European Patent Publication 0 242 236) which encodes phosphinothricin acetyl transferase (PAT). The enzyme PAT confers resistance to herbicidal glutamine synthetase inhibitors such as phosphinothricin. The *pat* gene in p35S/Ac is under the control of the 35S promoter from Cauliflower Mosaic Virus (Odell et al. (1985) *Nature* 313:810-812) and the 3' region of the nopaline synthase gene from the T-DNA of the Ti plasmid of *Agrobacterium tumefaciens*.

The particle bombardment method (Klein et al. (1987) *Nature* 327:70-73) may be used to transfer genes to the callus culture cells. According to this method, gold particles (1 μm in diameter) are coated with DNA using the following technique. Ten μg of plasmid DNAs are added to 50 μL of a suspension of gold particles (60 mg per mL). Calcium chloride (50 μL of a 2.5 M solution) and spermidine free base (20 μL of a 1.0 M solution) are added to the particles. The suspension is vortexed during the addition of these solutions. After 10 minutes, the tubes are briefly centrifuged (5 sec at 15,000 rpm) and the supernatant removed. The particles are resuspended in 200 μL of absolute ethanol, centrifuged again and the supernatant removed. The ethanol rinse is performed again and the particles resuspended in a final volume of 30 μL of ethanol. An aliquot (5 μL) of the DNA-coated gold particles can be placed in the center of a Kapton™ flying disc (Bio-Rad Labs). The particles are then accelerated into the corn tissue with a Biolistic™ PDS-1000/He (Bio-Rad Instruments, Hercules CA), using a helium pressure of 1000 psi, a gap distance of 0.5 cm and a flying distance of 1.0 cm.

For bombardment, the embryogenic tissue is placed on filter paper over agarose-solidified N6 medium. The tissue is arranged as a thin lawn and covered a circular area of about 5 cm in diameter. The petri dish containing the tissue can be placed in the chamber of the PDS-1000/He approximately 8 cm from the stopping screen. The air in the chamber is then evacuated to a vacuum of 28 inches of Hg. The macrocarrier is accelerated with a helium shock wave using a rupture membrane that bursts when the He pressure in the shock tube reaches 1000 psi.

Seven days after bombardment the tissue can be transferred to N6 medium that contains gluphosinate (2 mg per liter) and lacks casein or proline. The tissue continues to grow slowly on this medium. After an additional 2 weeks the tissue can be transferred to fresh N6 medium containing gluphosinate. After 6 weeks, areas of about 1 cm in diameter of actively growing callus can be identified on some of the plates containing the glufosinate-

supplemented medium. These calli may continue to grow when sub-cultured on the selective medium.

Plants can be regenerated from the transgenic callus by first transferring clusters of tissue to N6 medium supplemented with 0.2 mg per liter of 2,4-D. After two weeks the tissue can be transferred to regeneration medium (Fromm et al. (1990) *Bio/Technology 8*:833-839).

EXAMPLE 5

Expression of Chimeric Genes in Dicot Cells

A seed-specific expression cassette composed of the promoter and transcription terminator from the gene encoding the β subunit of the seed storage protein phaseolin from the bean *Phaseolus vulgaris* (Doyle et al. (1986) *J. Biol. Chem. 261*:9228-9238) can be used for expression of the instant polypeptides in transformed soybean. The phaseolin cassette includes about 500 nucleotides upstream (5') from the translation initiation codon and about 1650 nucleotides downstream (3') from the translation stop codon of phaseolin. Between the 5' and 3' regions are the unique restriction endonuclease sites Nco I (which includes the ATG translation initiation codon), SmaI, KpnI and XbaI. The entire cassette is flanked by Hind III sites.

The cDNA fragment of this gene may be generated by polymerase chain reaction (PCR) of the cDNA clone using appropriate oligonucleotide primers. Cloning sites can be incorporated into the oligonucleotides to provide proper orientation of the DNA fragment when inserted into the expression vector. Amplification is then performed as described above, and the isolated fragment is inserted into a pUC18 vector carrying the seed expression cassette.

Soybean embroys may then be transformed with the expression vector comprising sequences encoding the instant polypeptides. To induce somatic embryos, cotyledons, 3-5 mm in length dissected from surface sterilized, immature seeds of the soybean cultivar A2872, can be cultured in the light or dark at 26°C on an appropriate agar medium for 6-10 weeks. Somatic embryos which produce secondary embryos are then excised and placed into a suitable liquid medium. After repeated selection for clusters of somatic embryos which multiplied as early, globular staged embryos, the suspensions are maintained as described below.

Soybean embryogenic suspension cultures can maintained in 35 mL liquid media on a rotary shaker, 150 rpm, at 26°C with florescent lights on a 16:8 hour day/night schedule. Cultures are subcultured every two weeks by inoculating approximately 35 mg of tissue into 35 mL of liquid medium.

Soybean embryogenic suspension cultures may then be transformed by the method of particle gun bombardment (Klein et al. (1987) *Nature* (London) *327*:70-73, U.S. Patent

No. 4,945,050). A DuPont Biolistic™ PDS1000/HE instrument (helium retrofit) can be used for these transformations.

A selectable marker gene which can be used to facilitate soybean transformation is a chimeric gene composed of the 35S promoter from Cauliflower Mosaic Virus (Odell et al. (1985) *Nature 313*:810-812), the hygromycin phosphotransferase gene from plasmid pJR225 (from *E. coli*; Gritz et al.(1983) *Gene 25*:179-188) and the 3' region of the nopaline synthase gene from the T-DNA of the Ti plasmid of *Agrobacterium tumefaciens*. The seed expression cassette comprising the phaseolin 5' region, the fragment encoding the instant polypeptides and the phaseolin 3' region can be isolated as a restriction fragment. This fragment can then be inserted into a unique restriction site of the vector carrying the marker gene.

To 50 μL of a 60 mg/mL 1 μm gold particle suspension is added (in order): 5 μL DNA (1 μg/μL), 20 μl spermidine (0.1 M), and 50 μL $CaCl_2$ (2.5 M). The particle preparation is then agitated for three minutes, spun in a microfuge for 10 seconds and the supernatant removed. The DNA-coated particles are then washed once in 400 μL 70% ethanol and resuspended in 40 μL of anhydrous ethanol. The DNA/particle suspension can be sonicated three times for one second each. Five μL of the DNA-coated gold particles are then loaded on each macro carrier disk.

Approximately 300-400 mg of a two-week-old suspension culture is placed in an empty 60x15 mm petri dish and the residual liquid removed from the tissue with a pipette. For each transformation experiment, approximately 5-10 plates of tissue are normally bombarded. Membrane rupture pressure is set at 1100 psi and the chamber is evacuated to a vacuum of 28 inches mercury. The tissue is placed approximately 3.5 inches away from the retaining screen and bombarded three times. Following bombardment, the tissue can be divided in half and placed back into liquid and cultured as described above.

Five to seven days post bombardment, the liquid media may be exchanged with fresh media, and eleven to twelve days post bombardment with fresh media containing 50 mg/mL hygromycin. This selective media can be refreshed weekly. Seven to eight weeks post bombardment, green, transformed tissue may be observed growing from untransformed, necrotic embryogenic clusters. Isolated green tissue is removed and inoculated into individual flasks to generate new, clonally propagated, transformed embryogenic suspension cultures. Each new line may be treated as an independent transformation event. These suspensions can then be subcultured and maintained as clusters of immature embryos or regenerated into whole plants by maturation and germination of individual somatic embryos.

EXAMPLE 6

Expression of Chimeric Genes in Microbial Cells

The cDNAs encoding the instant polypeptides can be inserted into the T7 *E. coli* expression vector pBT430. This vector is a derivative of pET-3a (Rosenberg et al. (1987) *Gene 56*:125-135) which employs the bacteriophage T7 RNA polymerase/T7 promoter

system. Plasmid pBT430 was constructed by first destroying the EcoR I and Hind III sites in pET-3a at their original positions. An oligonucleotide adaptor containing EcoR I and Hind III sites was inserted at the BamH I site of pET-3a. This created pET-3aM with additional unique cloning sites for insertion of genes into the expression vector. Then, the Nde I site at the position of translation initiation was converted to an Nco I site using oligonucleotide-directed mutagenesis. The DNA sequence of pET-3aM in this region, 5'-CATATGG, was converted to 5'-CCCATGG in pBT430.

Plasmid DNA containing a cDNA may be appropriately digested to release a nucleic acid fragment encoding the protein. This fragment may then be purified on a 1% NuSieve GTG™ low melting agarose gel (FMC). Buffer and agarose contain 10 μg/ml ethidium bromide for visualization of the DNA fragment. The fragment can then be purified from the agarose gel by digestion with GELase™ (Epicentre Technologies) according to the manufacturer's instructions, ethanol precipitated, dried and resuspended in 20 μL of water. Appropriate oligonucleotide adapters may be ligated to the fragment using T4 DNA ligase (New England Biolabs, Beverly, MA). The fragment containing the ligated adapters can be purified from the excess adapters using low melting agarose as described above. The vector pBT430 is digested, dephosphorylated with alkaline phosphatase (NEB) and deproteinized with phenol/chloroform as described above. The prepared vector pBT430 and fragment can then be ligated at 16°C for 15 hours followed by transformation into DH5 electrocompetent cells (GIBCO BRL). Transformants can be selected on agar plates containing LB media and 100 μg/mL ampicillin. Transformants containing the gene encoding the instant polypeptides are then screened for the correct orientation with respect to the T7 promoter by restriction enzyme analysis.

For high level expression, a plasmid clone with the cDNA insert in the correct orientation relative to the T7 promoter can be transformed into *E. coli* strain BL21(DE3) (Studier et al. (1986) *J. Mol. Biol. 189*:113-130). Cultures are grown in LB medium containing ampicillin (100 mg/L) at 25°C. At an optical density at 600 nm of approximately 1, IPTG (isopropylthio-β-galactoside, the inducer) can be added to a final concentration of 0.4 mM and incubation can be continued for 3 h at 25°. Cells are then harvested by centrifugation and re-suspended in 50 μL of 50 mM Tris-HCl at pH 8.0 containing 0.1 mM DTT and 0.2 mM phenyl methylsulfonyl fluoride. A small amount of 1 mm glass beads can be added and the mixture sonicated 3 times for about 5 seconds each time with a microprobe sonicator. The mixture is centrifuged and the protein concentration of the supernatant determined. One μg of protein from the soluble fraction of the culture can be separated by SDS-polyacrylamide gel electrophoresis. Gels can be observed for protein bands migrating at the expected molecular weight.

CLAIMS

What is claimed is:

1. An isolated polynucleotide comprising a nucleotide sequence encoding a first polypeptide of at least 100 amino acids that has at least 80% identity based on the Clustal method of alignment when compared to a polypeptide selected from the group consisting of a *Momordica charantia* thioredoxin polypeptide of SEQ ID NO:2, a *Catalpa speciosa* thioredoxin polypeptide of SEQ ID NO:4, a soybean thioredoxin polypeptide of SEQ ID NO:6, a soybean thioredoxin polypeptide of SEQ ID NO:8 and a *Vernonia* thioredoxin polypeptide of SEQ ID NO:10.

2. An isolated polynucleotide comprising the complement of polynucleotide of Claim 1.

3. The isolated polynucleotide of Claim 1, wherein the nucleotide sequence comprises a nucleic acid sequence selected from the group consisting of SEQ ID NO:1, 3, 5, 7 and 9 that codes for the polypeptide selected from the group consisting of SEQ ID NO:2, 4, 6, 8 and 10.

4. The isolated polynucleotide of Claim 1 which is DNA.

5. The isolated polynucleotide of Claim 1 which is RNA.

6. A chimeric gene comprising the isolated polynucleotide of Claim 1 or Claim 2 operably linked to suitable regulatory sequences.

7. An isolated host cell comprising the chimeric gene of Claim 6.

8. An isolated host cell comprising an isolated polynucleotide of Claim 1.

9. The isolated host cell of Claim 8, wherein the host cell is yeast.

10. The isolated host cell of Claim 8, wherein the host cell is a bacterial cell.

11. The isolated host cell of Claim 8, wherein the host cell is a plant cell.

12. A virus comprising the isolated polynucleotide of Claim 1.

13. A process for producing an isolated host cell comprising the chimeric gene of claim 6, the process comprising either transforming or transfecting an isolated compatible host cell with the chimeric gene of Claim 6.

14. A thioredoxin polypeptide of at least 100 amino acids comprising at least 80% homology based on the Clustal method of alignment compared to a polypeptide selected from the group consisting of SEQ ID NO:2, 4, 6, 8 and 10.

15. A method of selecting an isolated polynucleotide that affects the level of expression of a thioredoxin polypeptide in a plant cell, the method comprising the steps of:

constructing an isolated polynucleotide comprising a nucleotide sequence of at least one of 30 contiguous nucleotides derived from a nucleotide sequence selected from the group consisting of SEQ ID NO:1, 3, 5, 7, 9 and the complement of such nucleotide sequences;

introducing the isolated polynucleotide into a plant cell;

measuring the level of thioredoxin polypeptide in the plant cell containing the polynucleotide; and

comparing the level of thioredoxin polypeptide in the plant cell containing the isolated polynucleotide with the level of thioredoxin polypeptide in a plant cell that does not contain the polynucleotide.

16. The method of Claim 15 wherein the isolated polynucleotide comprises a nucleotide sequence selected from the group consisting of SEQ ID NO:1, 3, 5, 7 and 9 that codes for the polypeptide selected from the group consisting of SEQ ID NO:2, 4, 6, 8 and 10.

17. The method of Claim 15 wherein the isolated polynucleotide is DNA.

18. The method of Claim 15 wherein the isolated polynucleotide is RNA.

19. The method of Claim 15 wherein the isolated polynucleotide is a chimeric gene comprising the nucleotide sequence operably linked to suitable regulatory sequences.

20. A method of selecting an isolated polynucleotide that affects the level of expression of thioredoxin polypeptide in a plant cell, the method comprising the steps of:

constructing the isolated polynucleotide of claim 1;

introducing the isolated polynucleotide into a plant cell;

measuring the level of thioredoxin polypeptide in the plant cell containing the polynucleotide; and

comparing the level of thioredoxin polypeptide in the plant cell containing the isolated polynucleotide with the level of thioredoxin polypeptide in a plant cell that does not contain the polynucleotide.

21. A method of obtaining a nucleic acid fragment encoding a substantial portion of a thioredoxin gene comprising the steps of:

synthesizing an oligonucleotide primer comprising a nucleotide sequence of at least one of 40 contiguous nucleotides derived from a nucleotide sequence selected from the group consisting of SEQ ID NO:1, 3, 5, 7, 9 and the complement of such nucleotide sequences; and

amplifying a nucleic acid sequence using the oligonucleotide primer.

22. A method of obtaining a nucleic acid fragment encoding all or a substantial portion of the amino acid sequence encoding a thioredoxin protein comprising the steps of:

probing a cDNA or genomic library with an isolated polynucleotide comprising a nucleotide sequence of at least one of 30 contiguous nucleotides derived from a nucleotide sequence selected from the group consisting of SEQ ID NO:1, 3, 5, 7, 9 and the complement of such nucleotide sequences;

identifying a DNA clone that hybridizes with the isolated polynucleotide;

isolating the identified DNA clone; and

sequencing the cDNA or genomic fragment that comprises the isolated DNA clone.

23. The isolated polynucleotide of Claim 1, wherein the first polypeptide is compared to the *Momordica charantia* thioredoxin polypeptide of SEQ ID NO:2.

24. The isolated polynucleotide of Claim 1, wherein the first polypeptide is compared to the *Catakoa speciosa* thioredoxin polypeptide of SEQ ID NO:4.

25. The isolated polynucleotide of Claim 1, wherein the first polypeptide is compared to the soybean thioredoxin polypeptide of SEQ ID NO:6.

26. The isolated polynucleotide of Claim 1, wherein the first polypeptide is compared to the soybean thioredoxin polypeptide of SEQ ID NO:8.

27. The isolated polynucleotide of Claim 1, wherein the first polypeptide is compared to the *Vernonia* thioredoxin polypeptide of SEQ ID NO:10.

SEQUENCE LISTING

<110> Allen, Steve
Lu, Albert
Thorpe, Cathy

<120> Thioredoxin H Homologs

<130> BB-1246

<140>
<141>

<150> 60/099,501
<151> September 8, 1998

<160> 13

<170> Microsoft Office 97

<210> 1
<211> 870
<212> DNA
<213> Momordica charantia

<400> 1
gcacgagggg gaagactttt tgcaggactg cgatttctgt ttacatcaat ggctgaagaa 60
gggcaagtga ttgcgtgtca taagatagat gaatgggagg gacaattagg aaaatggaag 120
gattctgaga aactggttgt ggtggatttt actgcttcct ggtgcgggcc atgccgggca 180
attgctccat atttcacaga attggctaag aataacccaa atgtcgcttt cctgaaagtc 240
gacgttgacg aattgaacag tgttgctagc aagtgggaga ttaatgcaat gccaacgttt 300
gttttcctga aaaaagggaa aataattgag aagatcgttg gtgctgataa agtggggctg 360
tcgaagaaaa tattagagct tagtggaact actcccgctg ctacttctac tgcttagaca 420
gtctgcttgg aggatgtgat ccctctggtg caatggtgat tccgctttgg agtttgatct 480
aattgtggat gaaactgtgt ctaaaagatg ttaattgttt ggccttttgg gttttcccct 540
ttttaagttt ggatcatgtg cgcacctctc agttgtgatt ctggtgctag aagcttcagg 600
tttcaatgtg gaataaatgg gggcacctgc tctgaaattg aatgacattt ttgcacactt 660
ttcattattc ttctgtaaga acttgaattc actgtttttt tttaatctaa ttcttcgtag 720
cagtacagtg agatgttctt tcagcttgtt tagcaacttc ttaatccctc tcctggcttt 780
tattttctta ttattggaat ggaacttaga agaatcgaag ttgttatgat ttgttaaaag 840
tatttgttgt taaaaaaaaa aaaaaaaaaa 870

<210> 2
<211> 122
<212> PRT
<213> Momordica charantia

<400> 2
Met Ala Glu Glu Gly Gln Val Ile Ala Cys His Lys Ile Asp Glu Trp
1 5 10 15

Glu Gly Gln Leu Gly Lys Trp Lys Asp Ser Glu Lys Leu Val Val Val
20 25 30

Asp Phe Thr Ala Ser Trp Cys Gly Pro Cys Arg Ala Ile Ala Pro Tyr
35 40 45

Phe Thr Glu Leu Ala Lys Asn Asn Pro Asn Val Ala Phe Leu Lys Val
50 55 60

```
Asp Val Asp Glu Leu Asn Ser Val Ala Ser Lys Trp Glu Ile Asn Ala
 65                  70                  75                  80

Met Pro Thr Phe Val Phe Leu Lys Lys Gly Lys Ile Ile Glu Lys Ile
                85                  90                  95

Val Gly Ala Asp Lys Val Gly Leu Ser Lys Lys Ile Leu Glu Leu Ser
            100                 105                 110

Gly Thr Thr Pro Ala Ala Thr Ser Thr Ala
        115                 120
```

<210> 3
<211> 574
<212> DNA
<213> Catalpa speciosa

<400> 3

```
gcacgagggc ataaatacct tgtaattggg gatttttcgg agtaaaaaag ggaaatcgga  60
aaatggcttc ttcagaagag ggacaagtga tcggttgcca ctccgtcgac gagtggaagg 120
agcagttcca gaagggtgtt gactctaaga aactggtggt aatagacttc acggcttcct 180
ggtgcggacc atgccgtttc attgctccaa tcttggctga tgatggccaag aagacacccc 240
atgtcatatt cctgaaagtc gacgtggatg aactcaagac tgttgctgag gaattcaaag 300
tggaggctat gccgaccttc gtgttcctca aggaagggaa agaagtggaa aggcttgtgg 360
gagcaaggaa ggaggaattg caggccacag ttgagaaaca tggcgctatc actgcttgat 420
gctgtttcaa tgtttagtta tgtaatatat gatgatgctt ggaataataa tgtcttaagt 480
tatccagatc gtatgtgact gacgtttctg ttgttatgtg gattgttatt gttaatgtaa 540
tgtaatggag tgtcttaaaa aaaaaaaaaa aaaa                              574
```

<210> 4
<211> 118
<212> PRT
<213> Catalpa speciosa

<400> 4

```
Met Ala Ser Ser Glu Glu Gly Gln Val Ile Gly Cys His Ser Val Asp
  1               5                  10                  15

Glu Trp Lys Glu Gln Phe Gln Lys Gly Val Asp Ser Lys Lys Leu Val
             20                  25                  30

Val Ile Asp Phe Thr Ala Ser Trp Cys Gly Pro Cys Arg Phe Ile Ala
         35                  40                  45

Pro Ile Leu Ala Glu Met Ala Lys Lys Thr Pro His Val Ile Phe Leu
     50                  55                  60

Lys Val Asp Val Asp Glu Leu Lys Thr Val Ala Glu Glu Phe Lys Val
 65                  70                  75                  80

Glu Ala Met Pro Thr Phe Val Phe Leu Lys Glu Gly Lys Glu Val Glu
                85                  90                  95

Arg Leu Val Gly Ala Arg Lys Glu Glu Leu Gln Ala Thr Val Glu Lys
            100                 105                 110

His Gly Ala Ile Thr Ala
        115
```

<210> 5
<211> 738
<212> DNA
<213> Glycine max

<400> 5
gcaccaggaa attctttagt tgtaactgac aaagttttct gagaaaataa ggattattat 60
tgagagaatg gctggctcat cggaagaggg acaagtcatt agctgccaca ccgttgaaga 120
atggaacgat caactccaga agggcaacga atccaagaaa ctcattgttg tggattttac 180
tgcttcttgg tgtggaccat gccgtttcat tgcaccattc ttggctgagc tggctaagaa 240
gttcacaagt gtcatattcc taaaggtgga tgtggacgaa ttaaagagtg tttctcaaga 300
ttgggctatt gaggctatgc ccacttttgt gtttgtgaaa gagggaacgc ttctggacaa 360
agtggtggga gcaaagaagg atgagctgca gcagaaaata cagaaacatg tggcttcagc 420
tagtgcttaa tctagctcac cttcagaaac tttatatatg cgctttcttt tcataatctt 480
gtactagact tatgttggta tttctgttat tgcaccaatc agcttttcaa aggtgatgac 540
tcctatcatc tatttctgaa tagtagtaac tggtcctttc ttccgtctta aataatagtg 600
gatggtgcta tatcatgaat cttaattaca tagaccttcc tgttttccct tttagtatta 660
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaacaataaa 720
aaaaacaaaa aaaaaaaa 738

<210> 6
<211> 120
<212> PRT
<213> Glycine max

<400> 6
Met Ala Gly Ser Ser Glu Glu Gly Gln Val Ile Ser Cys His Thr Val
1 5 10 15

Glu Glu Trp Asn Asp Gln Leu Gln Lys Gly Asn Glu Ser Lys Lys Leu
20 25 30

Ile Val Val Asp Phe Thr Ala Ser Trp Cys Gly Pro Cys Arg Phe Ile
35 40 45

Ala Pro Phe Leu Ala Glu Leu Ala Lys Lys Phe Thr Ser Val Ile Phe
50 55 60

Leu Lys Val Asp Val Asp Glu Leu Lys Ser Val Ser Gln Asp Trp Ala
65 70 75 80

Ile Glu Ala Met Pro Thr Phe Val Phe Val Lys Glu Gly Thr Leu Leu
85 90 95

Asp Lys Val Val Gly Ala Lys Lys Asp Glu Leu Gln Gln Lys Ile Gln
100 105 110

Lys His Val Ala Ser Ala Ser Ala
115 120

<210> 7
<211> 601
<212> DNA
<213> Glycine max

<400> 7
gcacgagctc tctctctcta gacttagatt ttgtgaatgg ctgaagtgga agagggacag 60
gtcatcggcg tccacaccgt tgatgagtgg aagctgcaac tccagaatgc aaaagactcc 120
aaaaaactga ttgtggtgga ttttactgct tcctggtgtg gtccatgccg ttttatggcc 180
ccagttcttg cagagattgc aaagaaaact cctgaattga tcttcctcaa agtggatgtg 240

gatgaagtga ggcctgttgc tgaggaatat tccattgagg ccatgccaac cttcctcttc 300
ttgaaagatg gcgagatcgt ggacaaggtg gttggtgcta gtaaggatga ccttcaagcc 360
accatagcca agcatgcatc tgctgttgct gctgcttctt cttcttgaag tgaagtatca 420
taatatgaaa gaagacaaag aataatgcat tttaatgttt tcaagtcagt ttggatgttt 480
tctctatgga cattgagttg gcagaacatc gagtgatgta taaaaataaa attgttgcat 540
tgtctttttt tcgtaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 600
a 601

<210> 8
<211> 123
<212> PRT
<213> Glycine max

<400> 8
Met Ala Glu Val Glu Glu Gly Gln Val Ile Gly Val His Thr Val Asp
1 5 10 15

Glu Trp Lys Leu Gln Leu Gln Asn Ala Lys Asp Ser Lys Lys Leu Ile
20 25 30

Val Val Asp Phe Thr Ala Ser Trp Cys Gly Pro Cys Arg Phe Met Ala
35 40 45

Pro Val Leu Ala Glu Ile Ala Lys Lys Thr Pro Glu Leu Ile Phe Leu
50 55 60

Lys Val Asp Val Asp Glu Val Arg Pro Val Ala Glu Glu Tyr Ser Ile
65 70 75 80

Glu Ala Met Pro Thr Phe Leu Phe Leu Lys Asp Gly Glu Ile Val Asp
85 90 95

Lys Val Val Gly Ala Ser Lys Asp Asp Leu Gln Ala Thr Ile Ala Lys
100 105 110

His Ala Ser Ala Val Ala Ala Ala Ser Ser Ser
115 120

<210> 9
<211> 614
<212> DNA
<213> Vernonia mespilifolia

<400> 9
gcacgaggct aaataccatt tgaaagctaa aaaaaaatct ttgaattagg ttttcttgaa 60
gaagtttgag aaaaaaaatg gcggaagaag gagttgtaac cggaatccac accgtcgacc 120
agtggaatga gcaacttgag aagcacaagg gaactgacaa attggtggtt gtggatttca 180
ccgcctcatg gtgtggtcct tgccgtgtga ttgcaccaat cttggctgat tttgctaaga 240
agatgcccca tgttaccttc cttaaggttg atgtggatga actcgagagc gttgctcagg 300
agtggtcagt ggaggcaatg ccgactttcc tgtttctcaa gggcggagtg aaagtggaca 360
aggttgtggg tgctaagaaa gacgaacttc atgcctgcat cgtcaagcat tctgctgcta 420
cagtttctgc ttaacgtact acataatatg attatcttat cagcaactta ttagtctctt 480
ttcggatgtg ttgttgattt gctttgtggt aaaaccttag attttgaata ttgtccttgt 540
aaccttgggt tataacttgc tctttcatct atatgcataa attgaagttg ctgtattaaa 600
aaaaaaaaaa aaaa 614

<210> 10
<211> 118
<212> PRT
<213> Vernonia mespilifolia

<400> 10
Met Ala Glu Glu Gly Val Val Thr Gly Ile His Thr Val Asp Gln Trp
1 5 10 15

Asn Glu Gln Leu Glu Lys His Lys Gly Thr Asp Lys Leu Val Val Val
20 25 30

Asp Phe Thr Ala Ser Trp Cys Gly Pro Cys Arg Val Ile Ala Pro Ile
35 40 45

Leu Ala Asp Phe Ala Lys Lys Met Pro His Val Thr Phe Leu Lys Val
50 55 60

Asp Val Asp Glu Leu Glu Ser Val Ala Gln Glu Trp Ser Val Glu Ala
65 70 75 80

Met Pro Thr Phe Leu Phe Leu Lys Gly Gly Val Lys Val Asp Lys Val
85 90 95

Val Gly Ala Lys Lys Asp Glu Leu His Ala Cys Ile Val Lys His Ser
100 105 110

Ala Ala Thr Val Ser Ala
115

<210> 11
<211> 114
<212> PRT
<213> Arabidopsis thaliana

<400> 11
Met Ala Ser Glu Glu Gly Gln Val Ile Ala Cys His Thr Val Glu Thr
1 5 10 15

Trp Asn Glu Gln Leu Gln Lys Ala Asn Glu Ser Lys Thr Leu Val Val
20 25 30

Val Asp Phe Thr Ala Ser Trp Cys Gly Pro Cys Arg Phe Ile Ala Pro
35 40 45

Phe Phe Ala Asp Leu Ala Lys Lys Leu Pro Asn Val Leu Phe Leu Lys
50 55 60

Val Asp Thr Asp Glu Leu Lys Ser Val Ala Ser Asp Trp Ala Ile Gln
65 70 75 80

Ala Met Pro Thr Phe Met Phe Leu Lys Glu Gly Lys Ile Leu Asp Lys
85 90 95

Val Val Gly Ala Lys Lys Asp Glu Leu Gln Ser Thr Ile Ala Lys His
100 105 110

Leu Ala

<210> 12
<211> 126
<212> PRT
<213> Nicotiana tabacum

<400> 12
Met Ala Ala Asn Asp Ala Thr Ser Ser Glu Glu Gly Gln Val Phe Gly
1 5 10 15

Cys His Lys Val Glu Glu Trp Asn Glu Tyr Phe Lys Lys Gly Val Glu
20 25 30

Thr Lys Lys Leu Val Val Val Asp Phe Thr Ala Ser Trp Cys Gly Pro
35 40 45

Cys Arg Phe Ile Ala Pro Ile Leu Ala Asp Ile Ala Lys Lys Met Pro
50 55 60

His Val Ile Phe Leu Lys Val Asp Val Asp Glu Leu Lys Thr Val Ser
65 70 75 80

Ala Glu Trp Ser Val Glu Ala Met Pro Thr Phe Val Phe Ile Lys Asp
85 90 95

Gly Lys Glu Val Asp Arg Val Val Gly Ala Lys Lys Glu Glu Leu Gln
100 105 110

Gln Thr Ile Val Lys His Ala Ala Pro Ala Thr Val Thr Ala
115 120 125

<210> 13
<211> 118
<212> PRT
<213> Ricinus communis

<400> 13
Met Ala Ala Glu Glu Gly Gln Val Ile Gly Cys His Thr Val Glu Ala
1 5 10 15

Trp Asn Glu Gln Leu Gln Lys Gly Asn Asp Thr Lys Gly Leu Ile Val
20 25 30

Val Asp Phe Thr Ala Ser Trp Cys Gly Pro Cys Arg Phe Ile Ala Pro
35 40 45

Phe Leu Ala Glu Leu Ala Lys Lys Leu Pro Asn Val Thr Phe Leu Lys
50 55 60

Val Asp Val Asp Glu Leu Lys Thr Val Ala His Glu Trp Ala Val Glu
65 70 75 80

Ser Met Pro Thr Phe Met Phe Leu Lys Glu Gly Lys Ile Met Asp Lys
85 90 95

Val Val Gly Ala Lys Lys Asp Glu Leu Gln Gln Thr Ile Ala Lys His
100 105 110

Met Ala Thr Ala Ser Thr
115

PATENT COOPERATION TREATY

From the
INTERNATIONAL PRELIMINARY EXAMINING AUTHORITY

To:

CHRISTENBURY, Lynne M.
E.I. DU PONT DE NEMOURS AND COMPANY
Legal/Patent Records Center
1007 Market Street
Wilmington, Delaware 19898
ETATS-UNIS D'AMERIQUE

TMR

PCT

NOTIFICATION OF TRANSMITTAL OF THE INTERNATIONAL PRELIMINARY EXAMINATION REPORT
(PCT Rule 71.1)

Date of mailing *(day/month/year)*	05.12.2000
Applicant's or agent's file reference: BB1246	IMPORTANT NOTIFICATION

International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/US99/20420	07/09/1999	08/09/1998

Applicant
E. I. DU PONT DE NEMOURS AND COMPANY et al.

1. The applicant is hereby notified that this International Preliminary Examining Authority transmits herewith the international preliminary examination report and its annexes, if any, established on the international application.

2. A copy of the report and its annexes, if any, is being transmitted to the International Bureau for communication to all the elected Offices.

3. Where required by any of the elected Offices, the International Bureau will prepare an English translation of the report (but not of any annexes) and will transmit such translation to those Offices.

4. **REMINDER**

The applicant must enter the national phase before each elected Office by performing certain acts (filing translations and paying national fees) within 30 months from the priority date (or later in some Offices) (Article 39(1)) (see also the reminder sent by the International Bureau with Form PCT/IB/301).

Where a translation of the international application must be furnished to an elected Office, that translation must contain a translation of any annexes to the international preliminary examination report. It is the applicant's responsibility to prepare and furnish such translation directly to each elected Office concerned.

For further details on the applicable time limits and requirements of the elected Offices, see Volume II of the PCT Applicant's Guide.

Name and mailing address of the IPEA/	Authorized officer
European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465	Vullo, C Tel.+49 89 2399-8061

REY NOTED

Form PCT/IPEA/416 (July 1992)

08 MAR 2001

PATENT COOPERATION TREATY

PCT

From the INTERNATIONAL BUREAU

NOTIFICATION OF THE RECORDING OF A CHANGE

(PCT Rule 92bis.1 and Administrative Instructions, Section 422)

To:

CHRISTENBURY, Lynne, M.
E.I. du Pont de Nemours and Company
Legal Patent Records Center
1007 Market Street
Wilmington, DE 19898
ETATS-UNIS D'AMERIQUE

Date of mailing (day/month/year)
02 May 2000 (02.05.00)

Applicant's or agent's file reference
BB1246

IMPORTANT NOTIFICATION

International application No.
PCT/US99/20420

International filing date (day/month/year)
07 September 1999 (07.09.99)

1. The following indications appeared on record concerning:

[ ] the applicant [ ] the inventor [X] the agent [ ] the common representative

Name and Address
MAJARIAN, William, R.
E.I. du Pont de Nemours and Company
Legal Patent Records Center
1007 Market Street
Wilmington, DE 19898
United States of America

State of Nationality:
State of Residence:
Telephone No.: 302-992-4926
Facsimile No.: 302-773-0164
Teleprinter No.:

2. The International Bureau hereby notifies the applicant that the following change has been recorded concerning:

[X] the person [ ] the name [ ] the address [ ] the nationality [ ] the residence

Name and Address
CHRISTENBURY, Lynne, M.
E.I. du Pont de Nemours and Company
Legal Patent Records Center
1007 Market Street
Wilmington, DE 19898
United States of America

State of Nationality:
State of Residence:
Telephone No.: 302-992-5481
Facsimile No.: 302-773-0164
Teleprinter No.:

3. Further observations, if necessary:

4. A copy of this notification has been sent to:

[X] the receiving Office
[ ] the International Searching Authority
[X] the International Preliminary Examining Authority
[ ] the designated Offices concerned
[X] the elected Offices concerned
[ ] other:

The International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20, Switzerland

Facsimile No.: (41-22) 740.14.35

Authorized officer
C. Cupello

Telephone No.: (41-22) 338.83.38

Form PCT/IB/306 (March 1994)

003257054

PATENT COOPERATION TREATY

From the INTERNATIONAL BUREAU

PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

To:

Assistant Commissioner for Patents
United States Patent and Trademark Office
Box PCT
Washington, D.C.20231
ETATS-UNIS D'AMERIQUE

in its capacity as elected Office

Date of mailing (day/month/year) 02 May 2000 (02.05.00)	
International application No. PCT/US99/20420	**Applicant's or agent's file reference** BB1246
International filing date (day/month/year) 07 September 1999 (07.09.99)	**Priority date** (day/month/year) 08 September 1998 (08.09.98)
Applicant ALLEN, Stephen, M. et al	

1. The designated Office is hereby notified of its election made:

[X] in the demand filed with the International Preliminary Examining Authority on:

22 March 2000 (22.03.00)

[ ] in a notice effecting later election filed with the International Bureau on:

2. The election [X] was

[ ] was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO **34, chemin des Colombettes** **1211 Geneva 20, Switzerland**	Authorized officer C. Cupello
Facsimile No.: (41-22) 740.14.35	Telephone No.: (41-22) 338.83.38